Express Mail Label No. EL251688920US

IN THE UNITED STATES PAPTENT AND TRADEMARK OFFICE

PROVISIONAL PATENT APPLICATION OF:

Ruoping **CHEN**, Chen W. **LIAW**, Derek T. **CHALMERS** and Dominic P. **BEHAN**

FOR:

# NON-ENDOGENOUS, CONSTITUTIVELY ACTIVATED HUMAN G PROTEIN-COUPLED ORPHAN RECEPTOR:TDAG8

ARENA PHARMACEUTICALS, INC.
6166 Nancy Ridge Drive
San Diego, CA 92121

NON-ENDOGENOUS CONSTITUTIVELY ACTIVATED
HUMAN G PROTEIN-COUPLED ORPHAN RECEPTOR: TDAG8

TABLE OF CONTENTS

# NON-ENDOGENOUS CONSTITUTIVELY ACTIVATED HUMAN G PROTEIN-COUPLED ORPHAN RECEPTOR: TDAG8

The benefit of commonly owned U.S. Serial Number 09/170,496, filed via Express Mail on October 13, 1998, is hereby claimed.

## FIELD OF THE INVENTION

The invention disclosed in this patent document relates to transmembrane receptors, more particularly to G protein-coupled receptors for which the endogenous ligand is unknown ("orphan GPCR receptors"), and most particularly, to a mutated (non-endogenous) version of an orphan GPCR receptor, TDAG8, that by virtue of the mutation is constitutively active.

## BACKGROUND OF THE INVENTION

Although a number of receptor classes exist in humans, by far the most abundant and therapeutically relevant is represented by the G protein-coupled receptor (GPCR or GPCRs) class. It is estimated that there are some 100,000 genes within the human genome, and of these, approximately 2% or 2,000 genes, are estimated to code for GPCRs. Of these, there are approximately 100 GPCRs for which the endogenous ligand that binds to the GPCR has been identified. Because of the significant time-lag that exists between the discovery of an endogenous GPCR and its endogenous ligand, it can be presumed that the remaining 1,900 GPCRs will be identified and characterized long before the endogenous ligands for these receptors are identified. Indeed, the rapidity by which the Human Genome Project is sequencing the 100,000 human genes indicates that the remaining human GPCRs

will be fully sequenced within the next few years. Nevertheless, and despite the efforts to sequence the human genome, it is still very unclear as to how scientists will be able to rapidly, effectively and efficiently exploit this information to improve and enhance the human condition. The present invention is geared towards this important objective.

Receptors, including GPCRs, for which the endogenous ligand has been identified are referred to as "known" receptors, while receptors for which the endogenous ligand has not been identified are referred to as "orphan" receptors. This distinction is not merely semantic, particularly in the case of GPCRs. GPCRs represent an important area for the development of pharmaceutical products: from approximately 20 of the 100 known GPCRs, 60% of all prescription pharmaceuticals have been developed. Thus, the orphan GPCRs are to the pharmaceutical industry what gold was to California in the late 19$^{th}$ century – an opportunity to drive growth, expansion, enhancement and development. A serious drawback exists, however, with orphan receptors relative to the discovery of novel therapeutics. This is because the traditional approach to the discovery and development of pharmaceuticals has required access to both the receptor *and* its endogenous ligand. Thus, heretofore, orphan GPCRs have presented the art with a tantalizing and undeveloped resource for the discovery of pharmaceuticals.

Under the traditional approach to the discovery of potential therapeutics, it is generally the case that the receptor is first identified. Before drug discovery efforts can be initiated, elaborate, time consuming and expensive procedures are typically put into place in order to identify, isolate and generate the receptor's endogenous ligand – this process can require from between three and ten years per receptor, at a cost of about $5million (U.S.) per receptor. These time and financial resources must be expended before the traditional

approach to drug discovery can commence. This is because traditional drug discovery techniques rely upon so-called "competitive binding assays" whereby putative therapeutic agents are "screened" against the receptor in an effort to discover compounds that either block the endogenous ligand from binding to the receptor ("antagonists"), or enhance or mimic the effects of the ligand binding to the receptor ("agonists"). The overall objective is to identify compounds that prevent cellular activation when the ligand binds to the receptor (the antagonists), or that enhance or increase cellular activity that would otherwise occur if the ligand was properly binding with the receptor (the agonists). Because the endogenous ligands for orphan GPCRs are by definition not identified, the ability to discover novel and unique therapeutics to these receptors using traditional drug discovery techniques is not possible. The present invention, as will be set forth in greater detail below, overcomes these and other severe limitations created by such traditional drug discovery techniques.

GPCRs share a common structural motif. All these receptors have seven sequences of between 22 to 24 hydrophobic amino acids that form seven alpha helices, each of which spans the membrane (each span is identified by number, *i.e.*, transmembrane-1 (TM-1), transmebrane-2 (TM-2), etc.). The transmembrane helices are joined by strands of amino acids between transmembrane-2 and transmembrane-3, transmembrane-4 and transmembrane-5, and transmembrane-6 and transmembrane-7 on the exterior, or "extracellular" side, of the cell membrane (these are referred to as "extracellular" regions 1, 2 and 3 (EC-1, EC-2 and EC-3), respectively). The transmembrane helices are also joined by strands of amino acids between transmembrane-1 and transmembrane-2, transmembrane-3 and transmembrane-4, and transmembrane-5 and transmembrane-6 on the interior, or "intracellular" side, of the cell membrane (these are referred to as "intracellular" regions 1, 2

and 3 (IC-1, IC-2 and IC-3), respectively). The "carboxy" ("C") terminus of the receptor lies in the intracellular space within the cell, and the "amino" ("N") terminus of the receptor lies in the extracellular space outside of the cell.

Generally, when an endogenous ligand binds with the receptor (often referred to as "activation" of the receptor), there is a change in the conformation of the intracellular region that allows for coupling between the intracellular region and an intracellular "G-protein." It has been reported that GPCRs are "promiscuous" with respect to G proteins, *i.e.,* that a GPCR can interact with more than one G protein. *See*, Kenakin, T., 43 *Life Sciences* 1095 (1988). Although other G proteins exist, currently, Gq, Gs, Gi, and Go are G proteins that have been identified. Endogenous ligand-activated GPCR coupling with the G-protein begins a signaling cascade process (referred to as "signal transduction"). Under normal conditions, signal transduction ultimately results in cellular activation or cellular inhibition. It is thought that the IC-3 loop as well as the carboxy terminus of the receptor interact with the G protein. A principal focus of this invention is directed to the transmembrane-6 (TM6) region and the intracellular-3 (IC3) region of the GPCR.

Under physiological conditions, GPCRs exist in the cell membrane in equilibrium between two different conformations: an "inactive" state and an "active" state. A receptor in an inactive state is unable to link to the intracellular signaling transduction pathway to produce a biological response. Changing the receptor conformation to the active state allows linkage to the transduction pathway (via the G-protein) and produces a biological response.

A receptor may be stabilized in an active state by an endogenous ligand or a compound such as a drug. Recent discoveries, including but not exclusively limited to

modifications to the amino acid sequence of the receptor, provide means other than endogenous ligands or drugs to promote and stabilize the receptor in the active state conformation. These means effectively stabilize the receptor in an active state by simulating the effect of an endogenous ligand binding to the receptor. Stabilization by such ligand-independent means is termed "constitutive receptor activation."

As noted above, the use of an orphan receptor for screening purposes has not been possible. This is because the traditional "dogma" regarding screening of compounds mandates that the ligand for the receptor be known. By definition, then, this approach has no applicability with respect to orphan receptors. Thus, by adhering to this dogmatic approach to the discovery of therapeutics, the art, in essence, has taught and has been taught to forsake the use of orphan receptors unless and until the endogenous ligand for the receptor is discovered. Given that there are an estimated 2,000 G protein coupled receptors, the majority of which are orphan receptors, such dogma castigates a creative, unique and distinct approach to the discovery of therapeutics.

The present invention relates to human T-cell death-associated gene receptor (TDAG8). TDAG8 was cloned and sequenced in 1998. Kyaw, H. et al, 17 DNA Cell Biol. 493 (1998); *see* Figure 1 of Kyaw for nucleic and deduced amino acid sequences. The endogenous ligand for TDAG8 is unknown. Thus, TDAG8 is an orphan GPCR having an open reading frame of 1,011 bp encoding a 337 amino acid protein. TDAG8 is reported to be homologous to the mouse TDAG8 and expressed in lymphoid tissues, including peripheral blood leukocytes, spleen, lymph nodes and thymus. TDAG8 is also reported to be localized to chromosome 14q31-32.1. *Id.*

As will be set forth and disclosed in greater detail below, utilization of a mutational cassette to modify the endogenous sequence of a human TDAG8 leads to a constitutively activated version of this receptor. This non-endogenous, constitutively activated version of human TDAG8 can be utilized, *inter alia*, for the screening of candidate compounds to directly identify compounds of, *e.g.*, therapeutic relevance.

## SUMMARY OF THE INVENTION

Disclosed herein is a mutated version of human TDAG8, which mutation leads to constitutive activation of this receptor.

## BRIEF DESCRIPTION OF THE DRAWINGS

**Figures 1A-1B** provides graphic results of comparative analysis of endogenous TDAG8 ("WT") versus non-endogenous, constitutively activated TDAG8 ("I225K") (control is designated "CMV") in 293 and 293T cells, respectively.

**Figure 2** is a schematic representation of a portion of the preferred 8XCRE-LUC Reporter plasmid construct utilized herein.

## DETAILED DESCRIPTION

The scientific literature that has evolved around receptors has adopted a number of terms to refer to ligands having various effects on receptors. For clarity and consistency, the following definitions will be used throughout this patent document. To the extent that these definitions conflict with other definitions for these terms, the following definitions shall control:

**AGONISTS** shall mean compounds that activate the intracellular response when they bind to the receptor, or enhance GTP binding to membranes.

**AMINO ACID ABBREVIATIONS** used herein are set below:

| | | |
|---|---|---|
| ALANINE | ALA | A |
| ARGININE | ARG | R |
| ASPARAGINE | ASN | N |
| ASPARTIC ACID | ASP | D |
| CYSTEINE | CYS | C |
| GLUTAMIC ACID | GLU | E |
| GLUTAMINE | GLN | Q |
| GLYCINE | GLY | G |
| HISTIDINE | HIS | H |
| ISOLEUCINE | ILE | I |
| LEUCINE | LEU | L |
| LYSINE | LYS | K |
| METHIONINE | MET | M |
| PHENYLALANINE | PHE | F |
| PROLINE | PRO | P |
| SERINE | SER | S |
| THREONINE | THR | T |
| TRYPTOPHAN | TRP | W |
| TYROSINE | TYR | Y |
| VALINE | VAL | V |

**PARTIAL AGONISTS** shall mean compounds which activate the intracellular response when they bind to the receptor to a lesser degree/extent than do agonists, or enhance GTP binding to membranes to a lesser degree/extent than do agonists

**ANTAGONIST** shall mean compounds that competitively bind to the receptor at the same site as the agonists but which do not activate the intracellular response initiated by

the active form of the receptor, and can thereby inhibit the intracellular responses by agonists or partial agonists. ANTAGONISTS do not diminish the baseline intracellular response in the absence of an agonist or partial agonist.

**CANDIDATE COMPOUND** shall mean a molecule (for example, and not limitation, a chemical compound) which is amenable to a screening technique. Preferably, the phrase "candidate compound" does not include compounds which were publicly known to be compounds selected from the group consisting of inverse agonist, agonist or antagonist to a receptor, as previously determined by an indirect identification process ("indirectly identified compound"); more preferably, not including an indirectly identified compound which has previously been determined to have therapeutic efficacy in at least one mammal; and, most preferably, not including an indirectly identified compound which has previously been determined to have therapeutic utility in humans.

**CODON** shall mean a grouping of three nucleotides (or equivalents to nucleotides) which generally comprise a nucleoside (adenosine (A), guanosine (G), cytidine (C), uridine (U) and thymidine (T)) coupled to a phosphate group and which, when translated, encodes an amino acid.

**COMPOUND EFFICACY** shall mean a measurement of the ability of a compound to inhibit or stimulate receptor functionality, as opposed to receptor binding affinity. A preferred means of detecting compound efficacy is via measurement of, *e.g.*, [$^{35}$S]GTPγS binding, as further disclosed in the Example section of this patent document.

**CONSTITUTIVELY ACTIVATED RECEPTOR** shall mean a receptor subject to constitutive receptor activation. In accordance with the invention disclosed herein, a non-

endogenous, human constitutively activated TDAG8 is one that has been sequence mutated relative to its endogenous sequence.

**CONSTITUTIVE RECEPTOR ACTIVATION** shall mean stabilization of a receptor in the active state by means other than binding of the receptor with its endogenous ligand or a chemical equivalent thereof. Preferably, a G protein-coupled receptor subjected to constitutive receptor activation in accordance with the invention disclosed herein evidences at least a 10% difference in response (increase or decrease, as the case may be) to the signal measured for constitutive activation as compared with the endogenous form of that GPCR, more preferably, about a 25% difference in such comparative response, and most preferably about a 50% difference in such comparative response. When used for the purposes of directly identifying candidate compounds, it is most preferred that the signal difference be at least about 50% such that there is a sufficient difference between the endogenous signal and the non-endogenous signal to differentiate between selected candidate compounds. In most instances, the "difference" will be an increase in signal; however, with respect to Gs-coupled GPCRs, the "difference" measured is preferably a decrease, as will be set forth in greater detail below.

**CONTACT** or **CONTACTING** shall mean bringing at least two moieties together, whether in an in vitro system or an in vivo system.

**DIRECTLY IDENTIFYING** or **DIRECTLY IDENTIFIED,** in relationship to the phrase "candidate compound", shall mean the screening of a candidate compound against a constitutively activated G protein-coupled receptor, and assessing the compound efficacy of such compound. This phrase is, under no circumstances, to be interpreted or

understood to be encompassed by or to encompass the phrase "indirectly identifying" or "indirectly identified."

**ENDOGENOUS** shall mean a material that is naturally produced by the genome of the species. ENDOGENOUS in reference to, for example and not limitation, GPCR, shall mean that which is naturally produced by a human, an insect, a plant, a bacterium, or a virus. By contrast, the term **NON-ENDOGENOUS** in this context shall mean that which is not naturally produced by the genome of a species. For example, and not limitation, a receptor which is not constitutively active in its endogenous form, but when mutated by using the cassettes disclosed herein and thereafter becomes constitutively active, is most preferably referred to herein as a "non-endogenous, constitutively activated receptor." Both terms can be utilized to describe both "in vivo" and "in vitro" systems. For example, and not limitation, in a screening approach, the endogenous or non-endogenous receptor may be in reference to an in vitro screening system whereby the receptor is expressed on the cell-surface of a mammalian cell. As a further example and not limitation, where the genome of a mammal has been manipulated to include a non-endogenous constitutively activated receptor, screening of a candidate compound by means of an in vivo system is viable.

**HOST CELL** shall mean a cell capable of having a Plasmid and/or Vector incorporated therein. In the case of a prokaryotic Host Cell, a Plasmid is typically replicated as an autonomous molecule as the Host Cell replicates (generally, the Plasmid is thereafter isolated for introduction into a eukaryotic Host Cell); in the case of a eukaryotic Host Cell, a Plasmid is integrated into the cellular DNA of the Host Cell such that when the eukaryotic Host Cell replicates, the Plasmid replicates. Preferably, for the purposes of the invention

disclosed herein, the Host Cell is eukaryotic, more preferably, mammalian, and most preferably selected from the group consisting of 293, 293T and COS-7 cells.

**INDIRECTLY IDENTIFYING** or **INDIRECTLY IDENTIFIED** means the traditional approach to the drug discovery process involving identification of an endogenous ligand specific for an endogenous receptor, screening of candidate compounds against the receptor for determination of those which interfere and/or compete with the ligand-receptor interaction, and assessing the efficacy of the compound for affecting at least one second messenger pathway associated with the activated receptor.

**INHIBIT** or **INHIBITING,** in relationship to the term "response" shall mean that a response is decreased or prevented in the presence of a compound as opposed to in the absence of the compound.

**INVERSE AGONISTS** shall mean compounds which bind to either the endogenous form of the receptor or to the constitutively activated form of the receptor, and which inhibit the baseline intracellular response initiated by the active form of the receptor below the normal base level of activity which is observed in the absence of agonists or partial agonists, or decrease GTP binding to membranes. Preferably, the baseline intracellular response is inhibited in the presence of the inverse agonist by at least 30%, more preferably by at least 50%, and most preferably by at least 75%, as compared with the baseline response in the absence of the inverse agonist.

**KNOWN RECEPTOR** shall mean an endogenous receptor for which the endogenous ligand specific for that receptor has been identified.

**LIGAND** shall mean an endogenous, naturally occurring molecule specific for an endogenous, naturally occurring receptor.

**MUTANT** or **MUTATION** in reference to an endogenous receptor's nucleic acid and/or amino acid sequence shall mean a specified change or changes to such endogenous sequences such that a mutated form of an endogenous, non-constitutively activated receptor evidences constitutive activation of the receptor. In terms of equivalents to specific sequences, a subsequent mutated form of a human receptor is considered to be equivalent to a first mutation of the human receptor if (a) the level of constitutive activation of the subsequent mutated form of the receptor is substantially the same as that evidenced by the first mutation of the receptor; and (b) the percent sequence (amino acid and/or nucleic acid) homology between the subsequent mutated form of the receptor and the first mutation of the receptor is at least about 80%, more preferably at least about 90% and most preferably at least 95%. Ideally, and owing to the fact that the most preferred mutation disclosed herein for achieving constitutive activation includes a single amino acid and/or codon change between the endogenous and the non-endogenous forms of the GPCR, the percent sequence homology should be at least 98%.

**ORPHAN RECEPTOR** shall mean an endogenous receptor for which the endogenous ligand specific for that receptor has not been identified or is not known.

**PHARMACEUTICAL COMPOSITION** shall mean a composition comprising at least one active ingredient, whereby the composition is amenable to investigation for a specified, efficacious outcome in a mammal (for example, and not limitation, a human). Those of ordinary skill in the art will understand and appreciate the techniques appropriate for determining whether an active ingredient has a desired efficacious outcome based upon the needs of the artisan.

**PLASMID** shall mean the combination of a Vector and cDNA. Generally, a Plasmid is introduced into a Host Cell for the purpose of replication and/or expression of the cDNA as a protein.

**STIMULATE** or **STIMULATING**, in relationship to the term "response" shall mean that a response is increased in the presence of a compound as opposed to in the absence of the compound.

**TRANSVERSE** or **TRANSVERSING**, in reference to either a defined nucleic acid sequence or a defined amino acid sequence, shall mean that the sequence is located within at least two different and defined regions. For example, in an amino acid sequence that is 10 amino acid moieties in length, where 3 of the 10 moieties are in the TM6 region of a GPCR and the remaining 7 moieties are in the IC3 region of the GPCR, the 10 amino acid moiety can be described as transversing the TM6 and IC3 regions of the GPCR.

**VECTOR** in reference to cDNA shall mean a circular DNA capable of incorporating at least one cDNA and capable of incorporation into a Host Cell.

The order of the following sections is set forth for presentational efficiency and is not intended, nor should be construed, as a limitation on the disclosure or the claims to follow.

## A. Introduction

The traditional study of receptors has always proceeded from the a priori assumption (historically based) that the endogenous ligand must first be identified before discovery could proceed to find antagonists and other molecules that could affect the receptor. Even in cases where an antagonist might have been known first, the search immediately extended to

looking for the endogenous ligand. This mode of thinking has persisted in receptor research even after the discovery of constitutively activated receptors. What has not been heretofore recognized is that it is the active state of the receptor that is most useful for discovering agonists, partial agonists, and inverse agonists of the receptor. For those diseases which result from an overly active receptor or an under-active receptor, what is desired in a therapeutic drug is a compound which acts to diminish the active state of a receptor or enhance the activity of the receptor, respectively, not necessarily a drug which is an antagonist to the endogenous ligand. This is because a compound that reduces or enhances the activity of the active receptor state need not bind at the same site as the endogenous ligand. Thus, as taught by a method of this invention, any search for therapeutic compounds should start by screening compounds against the ligand-independent active state.

Screening candidate compounds against non-endogenous, constitutively activated TDAG8 allows for the direct identification of candidate compounds which act at this cell surface receptor, without requiring any prior knowledge or use of the receptor's endogenous ligand. By determining areas within the body where the endogenous version of human TDAG8 are expressed and/or over-expressed, it is possible to determine related disease/disorder states which are associated with the expression and/or over-expression of the receptor; such an approach is disclosed in this patent document.

In its endogenous form, TDAG8 is not constitutively active, *i.e.*, TDAG8 signaling via G protein is ligand-dependent. Thus, it is not feasible to search directly for, *e.g.*, inverse agonists, to endogenous TDAG8. However, by the mutation approach disclosed in this patent document, TDAG8 can be constitutively activated such that screening of candidate compounds against the non-endogenous, constitutively activated TDAG8 orphan receptor

allows for the direct identification of candidate compounds as *e.g.*, inverse agonists. In the present invention, endogenous TDAG8 was mutated, such that the isoleucine residue at codon 225 was changed to a lysine residue, resulting in a non-endogenous, constitutively activated TDAG8. Although the non-endogenous amino acid at this position can be any of the amino acids (other than the endogenous isoleucine residue), most preferably, the non-endogenous amino acid is lysine.

Because there are only 20 naturally occurring amino acids (although the use of non-naturally occurring amino acids is also viable), selection of a particular non-endogenous amino acid for substitution at codon 225 is viable and allows for efficient selection of a non-endogenous amino acid that fits the needs of the investigator. However, as noted, the more preferred non-endogenous amino acids at codon 225 are lysine, hisitidine, arginine and alanine, with lysine being most preferred. Those of ordinary skill in the art are credited with the ability to readily determine proficient methods for changing the sequence of a codon to achieve a desired mutation.

**B. Disease/Disorder Identification and/or Selection**

As will be set forth in greater detail below, most preferably inverse agonists to the non-endogenous, constitutively activated TDAG8 receptor can be identified by the methodologies of this invention. Such inverse agonists are ideal candidates as lead compounds in drug discovery programs for treating diseases related to this receptor. Because of the ability to directly identify inverse agonists to the TDAG8 receptor, thereby allowing for the development of pharmaceutical compositions, a search for diseases and disorders associated with the TDAG8 receptor is relevant. For example, scanning both diseased and normal tissue samples for the presence of the TDAG8 receptor now becomes

more than an academic exercise or one which might be pursued along the path of identifying an endogenous ligand to TDAG8. Tissue scans can be conducted across a broad range of healthy and diseased tissues. Such tissue scans provide a preferred first step in associating a specific receptor with a disease and/or disorder.

Preferably, the DNA sequence of the TDAG8 receptor is used to make a probe for (a) dot-blot analysis against tissue-mRNA, and/or (b) RT-PCR identification of the expression of the receptor in tissue samples. The presence of a receptor in a tissue source, or a diseased tissue, or the presence of the receptor at elevated concentrations in diseased tissue compared to a normal tissue, can be preferably utilized to identify a correlation with a treatment regimen, including but not limited to, a disease associated with that disease. For example, TDAG8 is predominantly expressed in the lymphoid tissues, specifically the spleen, peripheral blood leukocytes and lymph nodes. Expression of TDAG8 has been reported to increase during activation of-induced death of T-cell hybridomas stimulated by glucocorticoids or anti-T-cell receptor antibodies (*see,* Choi J.W. et al. 168 Cell. Immunol. 78 (1996)). This report suggests that TDAG8 may play a role in immature thymocyte deletion and peripheral T-cell development. Thus, an inverse agonist to TDAG8 is intended to prevent the death of T-cells upon activation, which is an important role in the human immune system. Receptors can equally well be localized to regions of organs by this technique. Based on the known functions of the specific tissues to which the receptor is localized, the putative functional role of the receptor can be deduced.

## C. Screening of Candidate Compounds

### 1. Generic GPCR screening assay techniques

When a G protein receptor becomes constitutively active, it binds to a G protein (*e.g.*, Gq, Gs, Gi, Go) and stimulates the binding of GTP to the G protein. The G protein then acts as a GTPase and slowly hydrolyzes the GTP to GDP, whereby the receptor, under normal conditions, becomes deactivated. However, constitutively activated receptors continue to exchange GDP to GTP. A non-hydrolyzable analog of GTP, [$^{35}$S]GTPγS, can be used to monitor enhanced binding to membranes which express constitutively activated receptors. It is reported that [$^{35}$S]GTPγS can be used to monitor G protein coupling to membranes in the absence and presence of ligand. An example of this monitoring, among other examples well-known and available to those in the art, was reported by Traynor and Nahorski in 1995. The preferred use of this assay system is for initial screening of candidate compounds because the system is generically applicable to all G protein-coupled receptors regardless of the particular G protein that interacts with the intracellular domain of the receptor.

### 2. Specific GPCR screening assay techniques

Once candidate compounds are identified using the "generic" G protein-coupled receptor assay (*i.e.* an assay to select compounds that are agonists, partial agonists, or inverse agonists), further screening to confirm that the compounds have interacted at the receptor site is preferred. For example, a compound identified by the "generic" assay may not bind to the receptor, but may instead merely "uncouple" the G protein from the intracellular domain. In the case of TDAG8, it has been determined that this receptor couples the G protein Gs. Gs is known to activate the enzyme adenylyl cyclase, which is

necessary for catalyzing the conversion of ATP to cAMP. Thus, a non-endogenous, constitutively activated version of human TDAG8 would be expected to be associated with increased levels of cAMP. Thus, following direct identification of candidate compounds via a generic assay, an assay which is based upon responses mediated by the G protein associated with the receptor, *i.e.*, a "second messenger" assay, is preferably used to confirm or refine the direct identification results obtained from a generic assay.

Assays that detect cAMP can be utilized to determine if a candidate compound is e.g., an inverse agonist to a Gs-associated receptor (*i.e.*, such a compound would decrease the levels of cAMP) or a Gi-associated receptor (*i.e.*, such a candidate compound would increase the levels of cAMP). A variety of approaches known in the art for measuring cAMP can be utilized; a preferred approach relies upon the use of anti-cAMP antibodies in an ELISA-based format. Another type of assay, and most preferred, utilizes a whole cell second messenger reporter system assay. Promoters on genes drive the expression of the proteins that a particular gene encodes. Cyclic AMP drives gene expression by promoting the binding of a cAMP-responsive DNA binding protein or transcription factor (CREB) which then binds to the promoter at specific sites called cAMP response elements, which drives the expression of the gene. Preferably, reporter systems can be constructed which have multiple cAMP response elements before the reporter gene, *e.g.*, β-galactosidase or luciferase. Most preferably, the reporter system contains eight copies of the cAMP response elements (CREs) and an insertion of a rat somatostatin promoter region before the luciferase reporter gene. Thus, a constitutively activated Gs-linked receptor causes the accumulation of cAMP that then activates the gene and expression of the reporter protein. The reporter protein such as β-galactosidase or luciferase can then be detected using

standard biochemical assays (Chen W. *et al.* 226(2) *Anal. Biochem.* 349(1995)). Several reporter plasmids are known and available in the art for measuring a second messenger assay. It is considered well within the skilled artisan to determine an appropriate reporter plasmid for a particular gene expression based primarily upon the particular needs of the artisan.

The foregoing specific assay approach can, of course be utilized to initially directly identify candidate compounds, rather than by using the generic assay approach. Such a selection is primarily a matter of choice of the artisan.

**D. Medicinal Chemistry**

Generally, but not always, direct identification of candidate compounds is preferably conducted in conjunction with compounds generated via combinatorial chemistry techniques, whereby thousands of compounds are randomly prepared for such analysis. Generally, the results of such screening will be compounds having unique core structures; thereafter, these compounds are preferably subjected to additional chemical modification around a preferred core structure(s) to further enhance the medicinal properties thereof. Such techniques are known to those in the art and will not be addressed in detail in this patent document.

**E. Pharmaceutical compositions**

Candidate compounds selected for further development can be formulated into pharmaceutical compositions using techniques well known to those in the art. Suitable pharmaceutically-acceptable carriers are available to those in the art; for example, see Remington's Pharmaceutical Sciences, 16$^{th}$ Edition, 1980, Mack Publishing Co., (Oslo et al., eds.).

## F. Other Utility

Although a preferred use of the non-endogenous versions of human TDAG8 is for the direct identification of candidate compounds as inverse agonists, agonists or partial agonists (preferably for use as pharmaceutical agents), this version of human TDAG8 can also be utilized in research settings. For example, *in vitro* and *in vivo* systems incorporating TDAG8 can be utilized to further elucidate and understand the role(s) TDAG8 plays in the human condition, both normal and diseased, as well as understanding the role of constitutive activation as it applies to understanding the signaling cascade. A value in non-endogenous human TDAG8 is that its utility as a research tool is enhanced in that, because of its unique features, non-endogenous TDAG8 can be used to understand the role of TDAG8 in the human body before the endogenous ligand therefor is identified. Other uses of the disclosed receptors will become apparent to those in the art based upon, *inter alia*, a review of this patent document.

## EXAMPLES

The following examples are presented for purposes of elucidation, and not limitation, of the present invention. While specific nucleic acid and amino acid sequences are disclosed herein, those of ordinary skill in the art are credited with the ability to make minor modifications to these sequences while achieving the same or substantially similar results reported below.

### Example 1
### PREPARATION OF ENDOGENOUS, NON-CONSTITUTIVELY ACTIVATED TDAG8

PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each

primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of 94°C for 1 min, 56°C for 1min and 72 °C for 1 min and 20 sec. The 5' PCR primer contained a HindIII site with the following sequence:

5'-TGCAAGCTTAAAAAGGAAAAAATGAACAGC-3' (SEQ.ID.NO.:1)

and the 3' primer contained a BamHI site with the following sequence:

5'-TAAGGATCCCTTCCCTTCAAAACATCCTTG -3' (SEQ.ID.NO.:2).

The resulting 1.1 kb PCR fragment was digested with HindIII and BamHI and cloned into HindIII-BamHI site of pCMV expression vector. All 3 clones sequenced contained three potential polymorphisms involving changes of amino acid 43 from Pro to Ala, amino acid 97 from Lys to Asn and amino acid 130 from Ile to Phe. Nucleic acid (SEQ.ID.NO.:3) and amino acid (SEQ.ID.NO.:4) sequences for human TDAG8 were thereafter determined.

**Example 2**

**PREPARATION OF NON-ENDOGENOUS, CONSTITUTIVELY ACTIVATED TDAG8**

Preparation of the non-endogenous, constitutively activated human TDAG8 receptor was accomplished by creating a I225K mutation (see, SEQ.ID.NO.:5 for nucleic acid sequence, SEQ.ID.NO.:6 for amino acid sequence). Mutagenesis was performed using Transformer Site-Directed™ Mutagenesis Kit (Clontech) according to manufacturer's instructions. The two mutagenesis primers were utilized, a lysine mutagenesis oligonucleotide (SEQ.ID.NO.:7) and a selection marker oligonucleotide (SEQ.ID.NO.:8), which had the following sequences:

5'- GGAAAAGAAGAGAATCAAAAAACTACTTGTCAGCATC -3' (SEQ.ID.NO.: 7)

5'- CTCCTTCGGTCCTCCTATCGTTGTCAGAAGT -3' (SEQ.ID.NO.: 8),

respectively.

**Example 3**
**REPORTER-BASED ASSAY: CRE-LUC REPORTER ASSAY**

Although a variety of cells are available to the art for the expression of proteins, it is most preferred that mammalian cells be utilized. The primary reason for this is predicated upon practicalities, *i.e.*, utilization of, *e.g.*, yeast cells for the expression of a GPCR, while possible, introduces into the protocol a non-mammalian cell which may not (indeed, in the case of yeast, does not) include the receptor-coupling, genetic-mechanism and secretary pathways that have evolved for mammalian systems – thus, results obtained in non-mammalian cells, while of potential use, are not as preferred as that obtained from mammalian cells. Of the mammalian cells, COS-7, 293 and 293T cells are particularly preferred, although the specific mammalian cell utilized can be predicated upon the particular needs of the artisan.

293 and 293T cells were plated-out on 96 well plates at a density of $2 \times 10^4$ cells per well and were transfected using Lipofectamine Reagent (BRL) the following day according to manufacturer instructions. A DNA/lipid mixture was prepared for each 6-well transfection as follows: 260ng of plasmid DNA in 100μl of DMEM were gently mixed with 2μl of lipid in 100μl of DMEM (the 260ng of plasmid DNA consisted of 200ng of a 8xCRE-Luc reporter plasmid (*see* below and Figure 2 for a representation of a portion of the plasmid), 50ng of pCMV comprising endogenous receptor or non-endogenous receptor or pCMV alone, and 10ng of a GPRS expression plasmid (GPRS in pcDNA3 (Invitrogen)). The 8XCRE-Luc reporter plasmid was prepared as follows: vector SRIF-β-gal was obtained by cloning the rat somatostatin promoter (-71/+51) at BglV-HindIII site in the pβgal-Basic Vector (Clontech). Eight (8) copies of cAMP response element were obtained by PCR from an adenovirus template AdpCF126CCRE8

(*see* 7 Human Gene Therapy 1883 (1996)) and cloned into the SRIF-β-gal vector at the Kpn-BglV site, resulting in the 8xCRE-β-gal reporter vector. The 8xCRE-Luc reporter plasmid was generated by replacing the beta-galactosidase gene in the 8xCRE-β-gal reporter vector with the luciferase gene obtained from the pGL3-basic vector (Promega) at the HindIII-BamHI site. Following 30 min. incubation at room temperature, the DNA/lipid mixture was diluted with 400 μl of DMEM and 100μl of the diluted mixture was added to each well. 100 μl of DMEM with 10% FCS were added to each well after a 4hr incubation in a cell culture incubator. The following day the transfected cells were changed with 200 μl/well of DMEM with 10% FCS. Eight (8) hours later, the wells were changed to 100 μl /well of DMEM without phenol red, after one wash with PBS. Luciferase activity were measured the next day using the LucLite™ reporter gene assay kit (Packard) following manufacturer instructions and read on a 1450 MicroBeta™ scintillation and luminescence counter (Wallac). Results are summarized in Figures 1A and 1B.

Figure 1A represents an 81% increase in activity of the non-endogenous, constitutively active version of human TDAG8 (I225K) (14,440 relative light units) compared with that of the endogenous TDAG8 (WT) (2,715 relative light units) in 293 cells.

Figure 1B represents a 65% increase in activity of the non-endogenous, constitutively active version of human TDAG8 (I225K) (185,636 relative light units) compared with that of the endogenous GPR17 (WT) (65,681 relative light units) in 293T cells.

**Example 4**
**TISSUE DISTRIBUTION OF TDAG8**

Using a commercially available human-tissue dot-blot format, endogenous TDAG8 will be used to probe for a determination of the areas where such receptor is localized. The 1.1kb PCR fragment of Example 1 will be used as the probe: radiolabeled probe will be generated using this fragment and a Prime-It II™ Random Primer Labeling Kit (Stratagene, #300385), according to manufacturer's instructions. A human RNA Master Blot™ (Clontech, #7770-1) will be hybridized with TDAG8 radiolabeled probe and washed under stringent conditions according manufacturer's instructions. The blot will be exposed to Kodak BioMax Autoradiography film overnight at -80°C.

References cited throughout this patent document, unless otherwise indicated, are incorporated herein by reference. Modifications and extension of the disclosed inventions that are within the purview of the skilled artisan are encompassed within the above disclosure and the claims that follow.

Although a variety of expression vectors are available to those in the art, for purposes of utilization for both the endogenous and non-endogenous human TDAG8, it is most preferred that the vector utilized be pCMV. This vector was deposited with the American Type Culture Collection (ATCC) on October 13, 1998 (10801 University Blvd., Manassas, VA 20110-2209 USA) under the provisions of the Budapest Treaty for the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure. The DNA was tested by the ATCC and determined to be. The ATCC has assigned the following deposit number to pCMV: ATCC #203351.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Behan, Dominic P.
Chalmers, Derek T.
Liaw, Chen W.

(ii) TITLE OF INVENTION: Non-Endogenous, Constitutively Activated Human G Protein-Coupled Orphan Receptor: TDAG8

(iii) NUMBER OF SEQUENCES: 8

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Arena Pharmaceuticals, Inc.
(B) STREET: 6166 Nancy Ridge Drive
(C) CITY: San Diego
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 92122

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Burgoon, Richard P.
(B) REGISTRATION NUMBER: 34,787

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (619)453-7200
(B) TELEFAX: (619)453-7210

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TGCAAGCTTA AAAAGGAAAA AATGAACAGC 30

(3) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TAAGGATCCC TTCCCTTCAA AACATCCTTG 30

(4) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1014 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGAACAGCA CATGTATTGA AGAACAGCAT GACCTGGATC ACTATTTGTT TCCCATTGTT 60

TACATCTTTG TGATTATAGT CAGCATTCCA GCCAATATTG GATCTCTGTG TGTGTCTTTC 120

CTGCAACCCA AGAAGGAAAG TGAACTAGGA ATTTACCTCT TCAGTTTGTC ACTATCAGAT 180

TTACTCTATG CATTAACTCT CCCTTTATGG ATTGATTATA CTTGGAATAA AGACAACTGG 240

ACTTTCTCTC CTGCCTTGTG CAAAGGGAGT GCTTTTCTCA TGTACATGAA GTTTTACAGC 300

AGCACAGCAT TCCTCACCTG CATTGCCGTT GATCGGTATT TGGCTGTTGT CTACCCTTTG 360

AAGTTTTTTT TCCTAAGGAC AAGAAGAATT GCACTCATGG TCAGCCTGTC CATCTGGATA 420

TTGGAAACCA TCTTCAATGC TGTCATGTTG TGGGAAGATG AAACAGTTGT TGAATATTGC 480

GATGCCGAAA AGTCTAATTT TACTTTATGC TATGACAAAT ACCCTTTAGA GAAATGGCAA 540

ATCAACCTCA ACTTGTTCAG GACGTGTACA GGCTATGCAA TACCTTTGGT CACCATCCTG 600

ATCTGTAACC GGAAAGTCTA CCAAGCTGTG CGGCACAATA AAGCCACGGA AAACAAGGAA 660

AAGAAGAGAA TCATAAAACT ACTTGTCAGC ATCACAGTTA CTTTTGTCTT ATGCTTTACT 720

CCCTTTCATG TGATGTTGCT GATTCGCTGC ATTTTAGAGC ATGCTGTGAA CTTCGAAGAC 780

CACAGCAATT CTGGGAAGCG AACTTACACA ATGTATAGAA TCACGGTTGC ATTAACAAGT 840

TTAAATTGTG TTGCTGATCC AATTCTGTAC TGTTTTGTTA CCGAAACAGG AAGATATGAT 900

ATGTGGAATA TATTAAAATT CTGCACTGGG AGGTGTAATA CATCACAAAG ACAAAGAAAA 960

CGCATACTTT CTGTGTCTAC AAAAGATACT ATGGAATTAG AGGTCCTTGA GTAG 1014

(5) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 337 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Asn Ser Thr Cys Ile Glu Glu Gln His Asp Leu Asp His Tyr Leu
1 5 10 15

Phe Pro Ile Val Tyr Ile Phe Val Ile Ile Val Ser Ile Pro Ala Asn
20 25 30

Ile Gly Ser Leu Cys Val Ser Phe Leu Gln Pro Lys Lys Glu Ser Glu
35 40 45

Leu Gly Ile Tyr Leu Phe Ser Leu Ser Leu Ser Asp Leu Leu Tyr Ala
50 55 60

Leu Thr Leu Pro Leu Trp Ile Asp Tyr Thr Trp Asn Lys Asp Asn Trp
65 70 75 80

Thr Phe Ser Pro Ala Leu Cys Lys Gly Ser Ala Phe Leu Met Tyr Met
85 90 95

Lys Phe Tyr Ser Ser Thr Ala Phe Leu Thr Cys Ile Ala Val Asp Arg
100 105 110

Tyr Leu Ala Val Val Tyr Pro Leu Lys Phe Phe Phe Leu Arg Thr Arg
115 120 125

Arg Ile Ala Leu Met Val Ser Leu Ser Ile Trp Ile Leu Glu Thr Ile
130 135 140

Phe Asn Ala Val Met Leu Trp Glu Asp Glu Thr Val Val Glu Tyr Cys
145 150 155 160

Asp Ala Glu Lys Ser Asn Phe Thr Leu Cys Tyr Asp Lys Tyr Pro Leu
165 170 175

Glu Lys Trp Gln Ile Asn Leu Asn Leu Phe Arg Thr Cys Thr Gly Tyr
180 185 190

Ala Ile Pro Leu Val Thr Ile Leu Ile Cys Asn Arg Lys Val Tyr Gln
195 200 205

```
Ala Val Arg His Asn Lys Ala Thr Glu Asn Lys Glu Lys Lys Arg Ile
    210                 215                 220

Ile Lys Leu Leu Val Ser Ile Thr Val Thr Phe Val Leu Cys Phe Thr
225                 230                 235                 240

Pro Phe His Val Met Leu Leu Ile Arg Cys Ile Leu Glu His Ala Val
                245                 250                 255

Asn Phe Glu Asp His Ser Asn Ser Gly Lys Arg Thr Tyr Thr Met Tyr
            260                 265                 270

Arg Ile Thr Val Ala Leu Thr Ser Leu Asn Cys Val Ala Asp Pro Ile
        275                 280                 285

Leu Tyr Cys Phe Val Thr Glu Thr Gly Arg Tyr Asp Met Trp Asn Ile
    290                 295                 300

Leu Lys Phe Cys Thr Gly Arg Cys Asn Thr Ser Gln Arg Gln Arg Lys
305                 310                 315                 320

Arg Ile Leu Ser Val Ser Thr Lys Asp Thr Met Glu Leu Glu Val Leu
                325                 330                 335

Glu
```

(6) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1014 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
ATGAACAGCA CATGTATTGA AGAACAGCAT GACCTGGATC ACTATTTGTT TCCCATTGTT    60

TACATCTTTG TGATTATAGT CAGCATTCCA GCCAATATTG GATCTCTGTG TGTGTCTTTC   120

CTGCAAGCAA AGAAGGAAAG TGAACTAGGA ATTTACCTCT TCAGTTTGTC ACTATCAGAT   180

TTACTCTATG CATTAACTCT CCCTTTATGG ATTGATTATA CTTGGAATAA AGACAACTGG   240

ACTTTCTCTC CTGCCTTGTG CAAAGGGAGT GCTTTTCTCA TGTACATGAA TTTTTACAGC   300

AGCACAGCAT TCCTCACCTG CATTGCCGTT GATCGGTATT TGGCTGTTGT CTACCCTTTG   360

AAGTTTTTTT TCCTAAGGAC AAGAAGATTT GCACTCATGG TCAGCCTGTC CATCTGGATA   420

TTGGAAACCA TCTTCAATGC TGTCATGTTG TGGGAAGATG AAACAGTTGT TGAATATTGC   480
```

GATGCCGAAA AGTCTAATTT TACTTTATGC TATGACAAAT ACCCTTTAGA GAAATGGCAA 540

ATCAACCTCA ACTTGTTCAG GACGTGTACA GGCTATGCAA TACCTTTGGT CACCATCCTG 600

ATCTGTAACC GGAAAGTCTA CCAAGCTGTG CGGCACAATA AAGCCACGGA AAACAAGGAA 660

AAGAAGAGAA TCAAAAAACT ACTTGTCAGC ATCACAGTTA CTTTTGTCTT ATGCTTTACT 720

CCCTTTCATG TGATGTTGCT GATTCGCTGC ATTTTAGAGC ATGCTGTGAA CTTCGAAGAC 780

CACAGCAATT CTGGGAAGCG AACTTACACA ATGTATAGAA TCACGGTTGC ATTAACAAGT 840

TTAAATTGTG TTGCTGATCC AATTCTGTAC TGTTTTGTTA CCGAAACAGG AAGATATGAT 900

ATGTGGAATA TATTAAAATT CTGCACTGGG AGGTGTAATA CATCACAAAG ACAAAGAAAA 960

CGCATACTTT CTGTGTCTAC AAAAGATACT ATGGAATTAG AGGTCCTTGA GTAG 1014

(7) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 337 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asn Ser Thr Cys Ile Glu Glu Gln His Asp Leu Asp His Tyr Leu
1 5 10 15

Phe Pro Ile Val Tyr Ile Phe Val Ile Ile Val Ser Ile Pro Ala Asn
20 25 30

Ile Gly Ser Leu Cys Val Ser Phe Leu Gln Ala Lys Lys Glu Ser Glu
35 40 45

Leu Gly Ile Tyr Leu Phe Ser Leu Ser Leu Ser Asp Leu Leu Tyr Ala
50 55 60

Leu Thr Leu Pro Leu Trp Ile Asp Tyr Thr Trp Asn Lys Asp Asn Trp
65 70 75 80

Thr Phe Ser Pro Ala Leu Cys Lys Gly Ser Ala Phe Leu Met Tyr Met
85 90 95

Asn Phe Tyr Ser Ser Thr Ala Phe Leu Thr Cys Ile Ala Val Asp Arg
100 105 110

Tyr Leu Ala Val Val Tyr Pro Leu Lys Phe Phe Phe Leu Arg Thr Arg
115 120 125

Arg Phe Ala Leu Met Val Ser Leu Ser Ile Trp Ile Leu Glu Thr Ile
130 135 140

```
Phe Asn Ala Val Met Leu Trp Glu Asp Glu Thr Val Val Glu Tyr Cys
145                 150                 155                 160

Asp Ala Glu Lys Ser Asn Phe Thr Leu Cys Tyr Asp Lys Tyr Pro Leu
                165                 170                 175

Glu Lys Trp Gln Ile Asn Leu Asn Leu Phe Arg Thr Cys Thr Gly Tyr
            180                 185                 190

Ala Ile Pro Leu Val Thr Ile Leu Ile Cys Asn Arg Lys Val Tyr Gln
        195                 200                 205

Ala Val Arg His Asn Lys Ala Thr Glu Asn Lys Glu Lys Lys Arg Ile
    210                 215                 220

Lys Lys Leu Leu Val Ser Ile Thr Val Thr Phe Val Leu Cys Phe Thr
225                 230                 235                 240

Pro Phe His Val Met Leu Leu Ile Arg Cys Ile Leu Glu His Ala Val
                245                 250                 255

Asn Phe Glu Asp His Ser Asn Ser Gly Lys Arg Thr Tyr Thr Met Tyr
            260                 265                 270

Arg Ile Thr Val Ala Leu Thr Ser Leu Asn Cys Val Ala Asp Pro Ile
        275                 280                 285

Leu Tyr Cys Phe Val Thr Glu Thr Gly Arg Tyr Asp Met Trp Asn Ile
    290                 295                 300

Leu Lys Phe Cys Thr Gly Arg Cys Asn Thr Ser Gln Arg Gln Arg Lys
305                 310                 315                 320

Arg Ile Leu Ser Val Ser Thr Lys Asp Thr Met Glu Leu Glu Val Leu
                325                 330                 335

Glu
```

(8) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGAAAAGAAG AGAATCAAAA AACTACTTGT CAGCATC 37

(9) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTCCTTCGGT CCTCCTATCG TTGTCAGAAG T 31

## ABSTRACT

The invention disclosed in this patent document relates to transmembrane receptors, more particularly to G protein-coupled receptors for which the endogenous ligand is unknown ("orphan GPCR receptors"), and most particularly to a mutated (non-endogenous) version of human TDAG8, with such mutated version being constitutively active.

FIGURE 1A

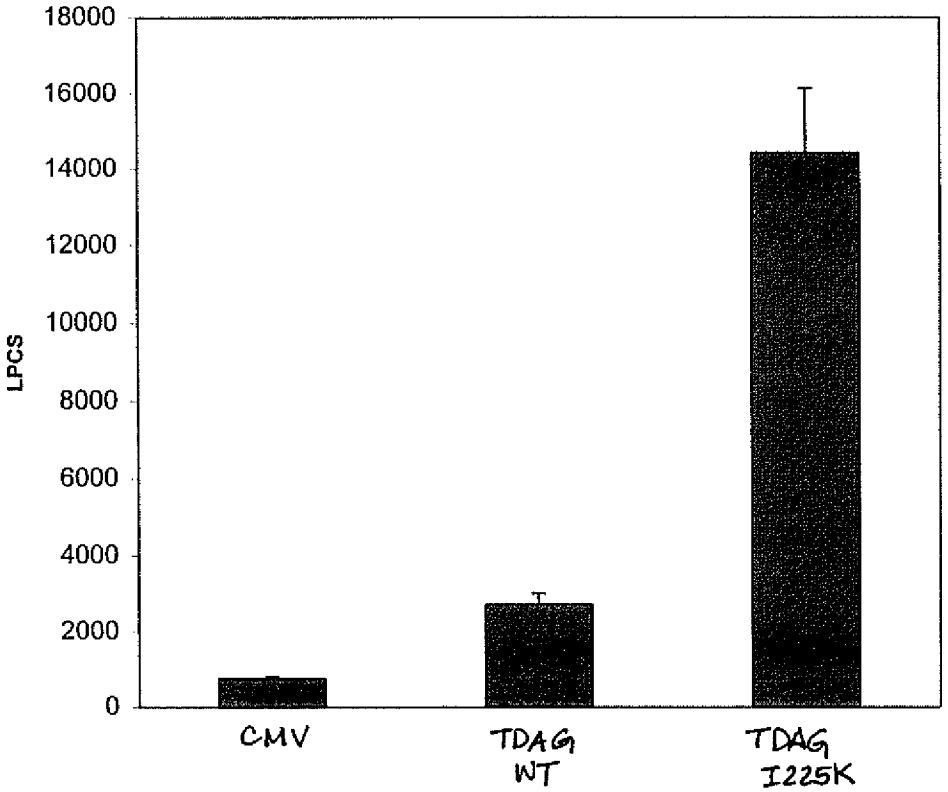
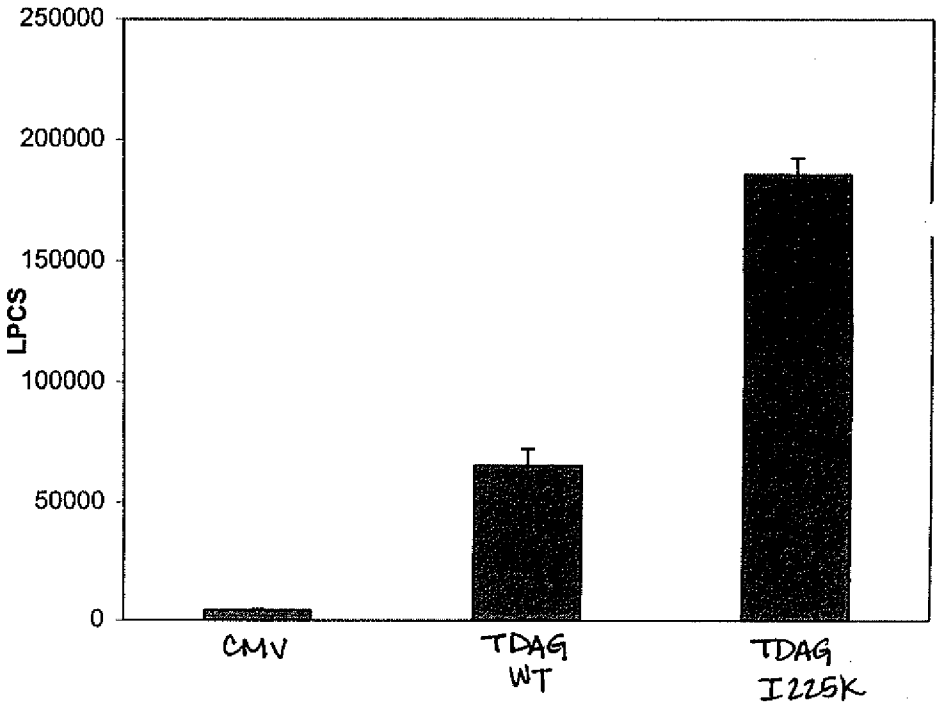



FIGURE 1B



# 8 x CRE-luc Reporter

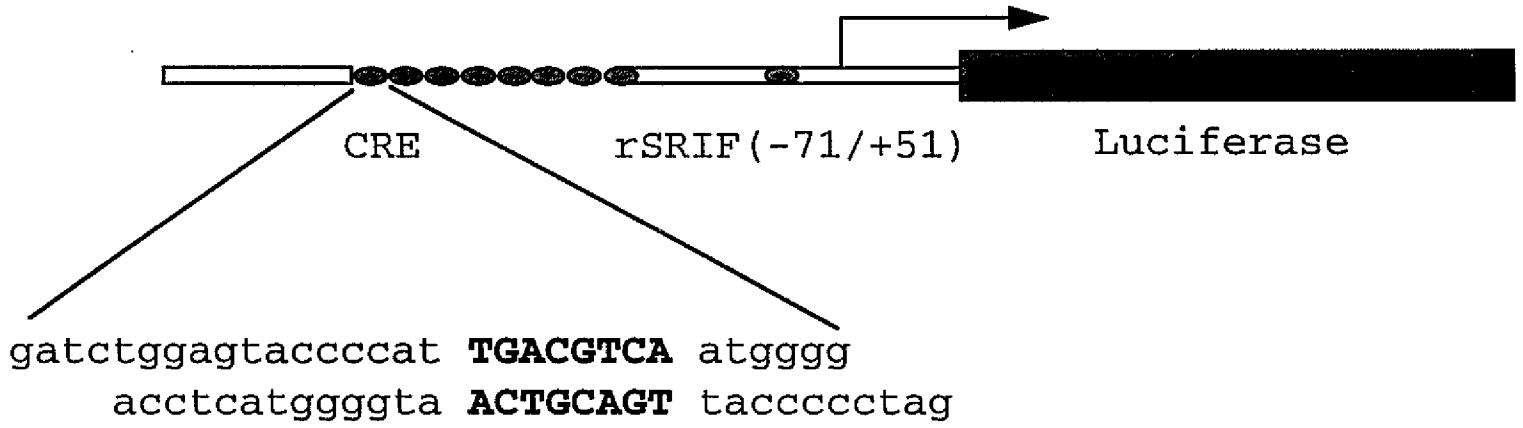



FIGURE 2